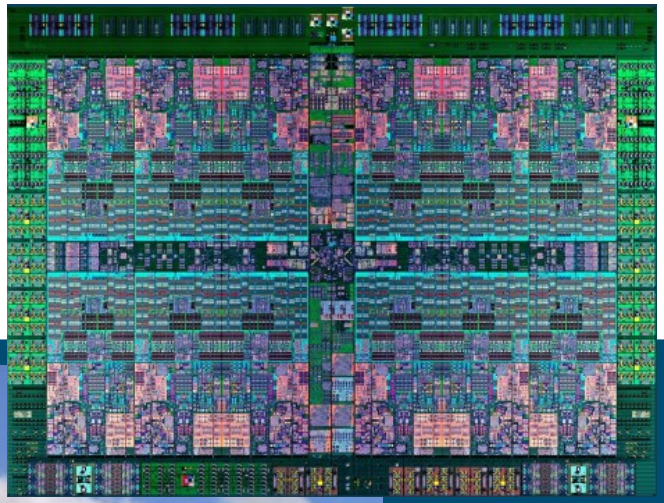
Mitglied der Helmholtz-Gemeinschaft

OpenPOWER™

JÜLICH
FORSCHUNGSZENTRUM



Von Antreibern und Beschleunigern des HPC

D. Pleiter | Jülich | 16 December 2014

Ein Dementi vorweg

kann. Das Forschungszentrum Jülich etwa, ebenfalls ein Mitglied des OpenPower-Konsortiums, wird sicherlich bei IBM bleiben und vom abgekündigten Blue Gene auf Power8 umschwenken. Mit IBM und Nvidia hat Jülich nun zur Einstimmung ein Power Acceleration and Design Center gegründet.

[c't, Nr. 25/2014, 15.11.2014]

- **Ja:** Das FZJ ist seit März Mitglieder der OpenPOWER Foundation und plant das POWER Acceleration and Design Center
- **Nein:** Kein Lock-in: wir legen uns heute noch nicht auf einen Hersteller fest
- **Aber:** **Wir denken natürlich viel über zukünftige Architekturen nach ...**

OpenPOWER Foundation

History

- Announced in September 2013
- Established in December 2013 as an open, not-for-profit technical membership group
- November 2014: >60 institutions joined

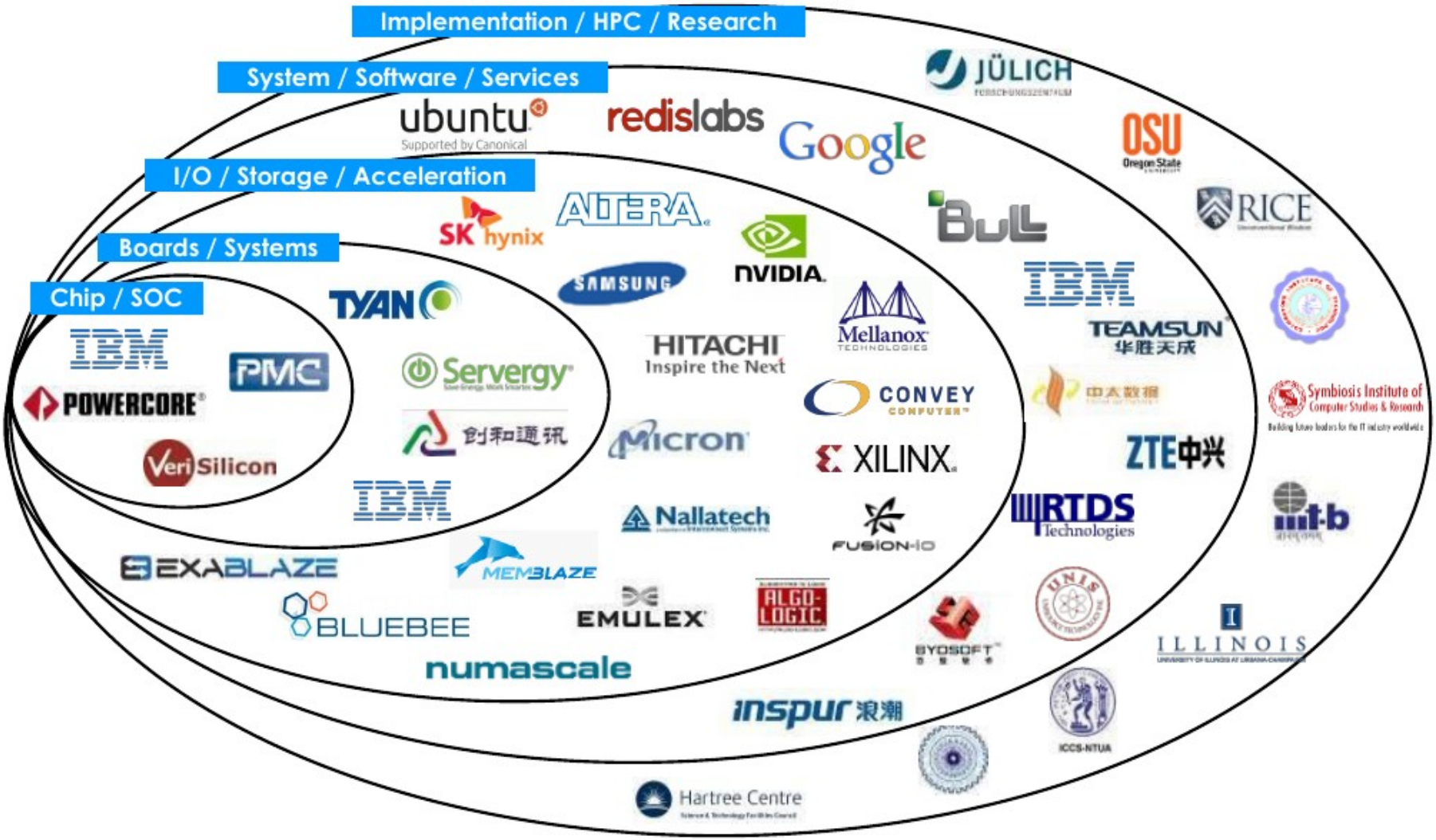
Mission statement

- Create an open ecosystem, using the POWER Architecture to share expertise, investment, and server-class intellectual property to serve the evolving needs of customers and industry

Organisation

- Board of directors + Technical Steering Committee
- Work groups

OpenPOWER Foundation (cont.)



Long-term HPC Trends: Top500 List

Performance metric

- Floating-point operations per time unit while solving a dense linear set of equations

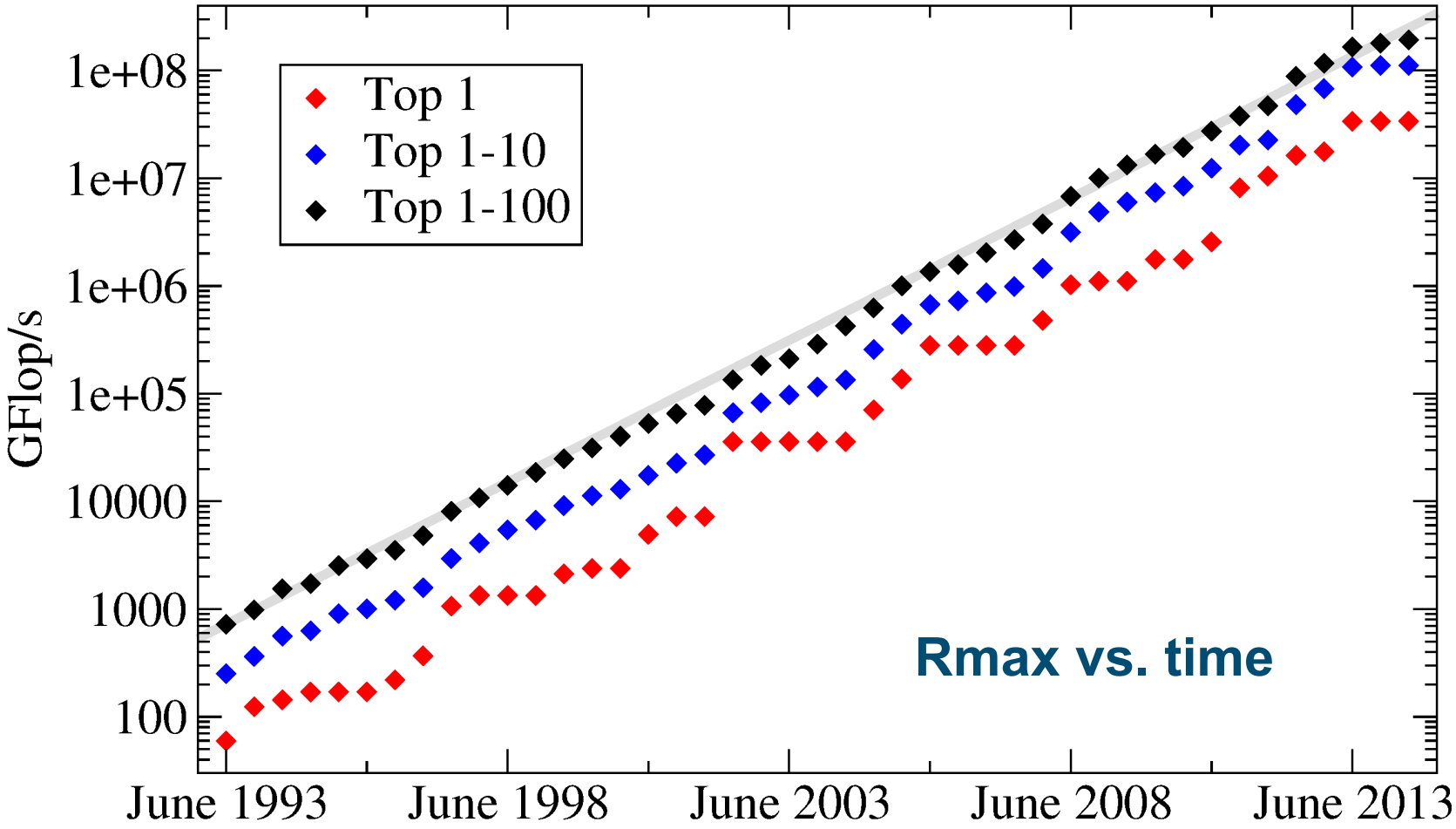
Criticism

- Workload not representative
- Problem size can be freely tuned

Positive aspects

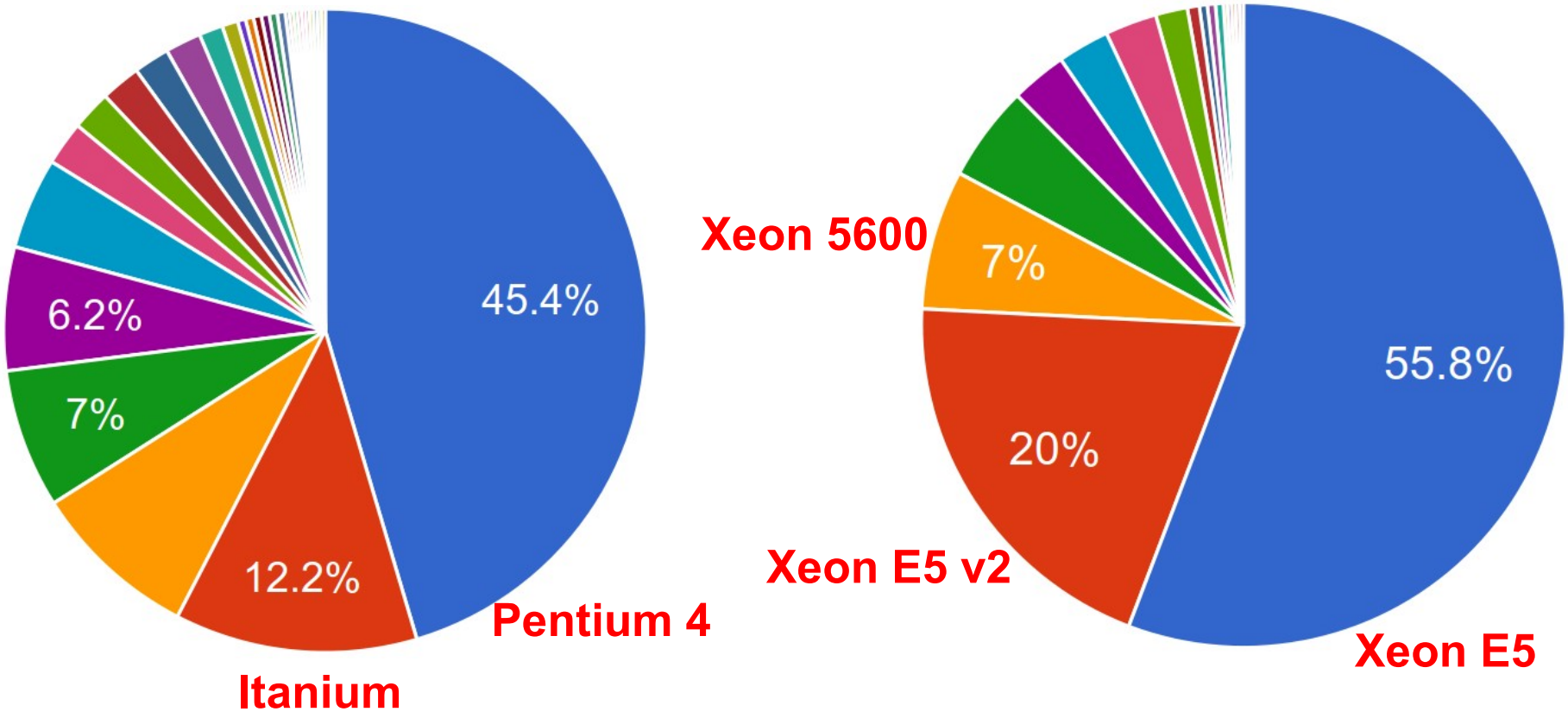
- Reasonable well defined basis for comparison
- Allows for long-term comparison
  - First list published in June 1993

Top500 Performance Trends



Top500 Processor Architecture Trends

System share:



Why OpenPOWER? A Customer View

Increasing share of Top500 are based on CPUs from single vendor

- Pure market observation, no statement about technology

Lack of competition

- Usually higher prices
- Less incentive for innovations

Need for promoting alternative technologies

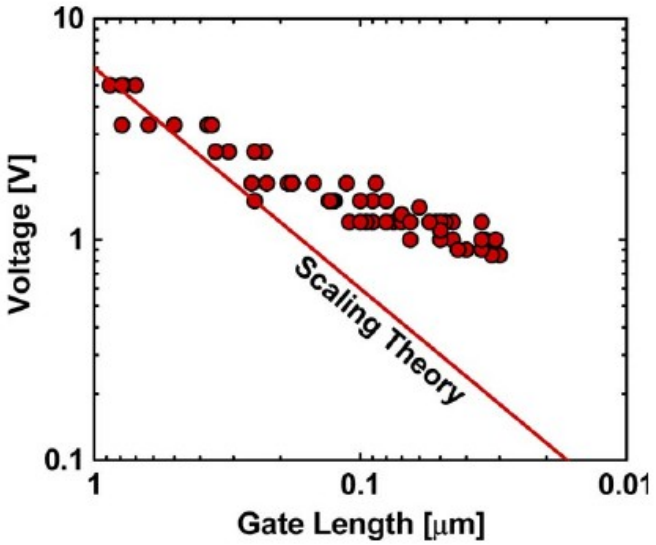
- OpenPOWER
- ARM

Key Technology Constraints

Dennard Scaling for MOSFET transistors

- Allowed for change of following parameters such that electric fields are roughly constant:
  - Transistor density
  - Switching speed
  - Supply voltage

[L. Chang et al., 2010]



Breakdown of Dennard Scaling

- Broken since around 2005 due to the end of voltage scaling
- Scaling of switching speed prohibitive due to power consumption

☞ **More performance = more parallelism**

Technology Path: More Parallel Processors

Processor parallelism

- Micro-architecture level:
  - Data-parallel instructions (SIMD)
  - Number of instruction pipelines
- Processor level: multi-core

Example: JUROPA Cluster at JSC

	JUROPA-2	JUROPA-4
SIMD width	2x64 bit	4x64 bit
No. of SIMD pipelines	1	2
Core/processor	4	12
Flop/cycle/processor	16	192
Core clock frequency [GHz]	2.93	2.5

Even more Parallel ”Accelerators” ...

Competing technologies

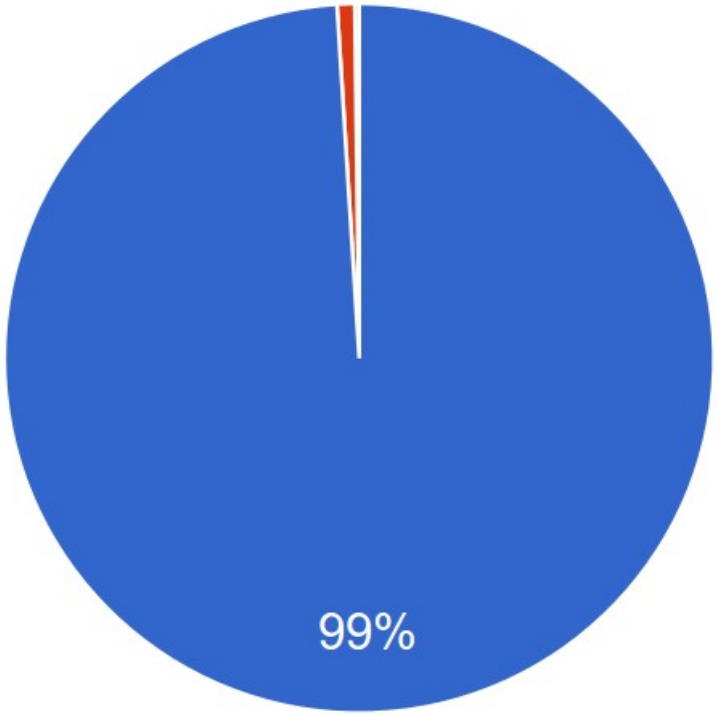
- Graphics processing units (GPU)
- Xeon Phi

Processor level parallelism

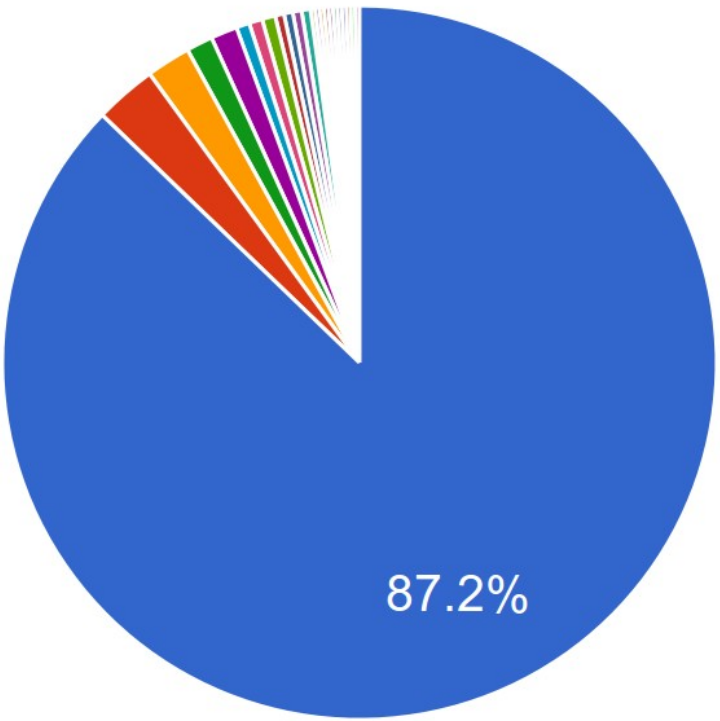
	NVIDIA K40	Intel Xeon Phi 7120D
Flop/cycle/processor	1920	976
Core clock frequency [GHz]	0.75	1.24

Top500 Trends on Accelerated Architectures

System share:



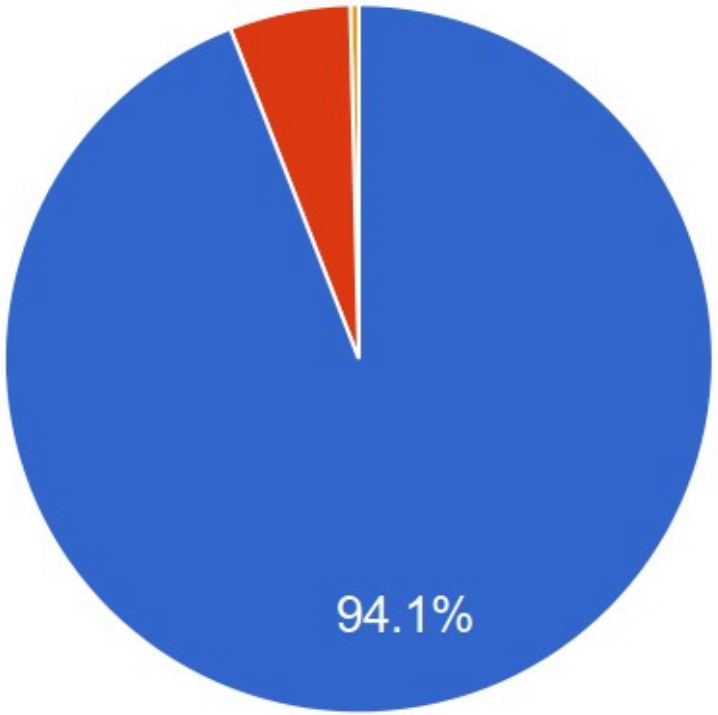
June 2009



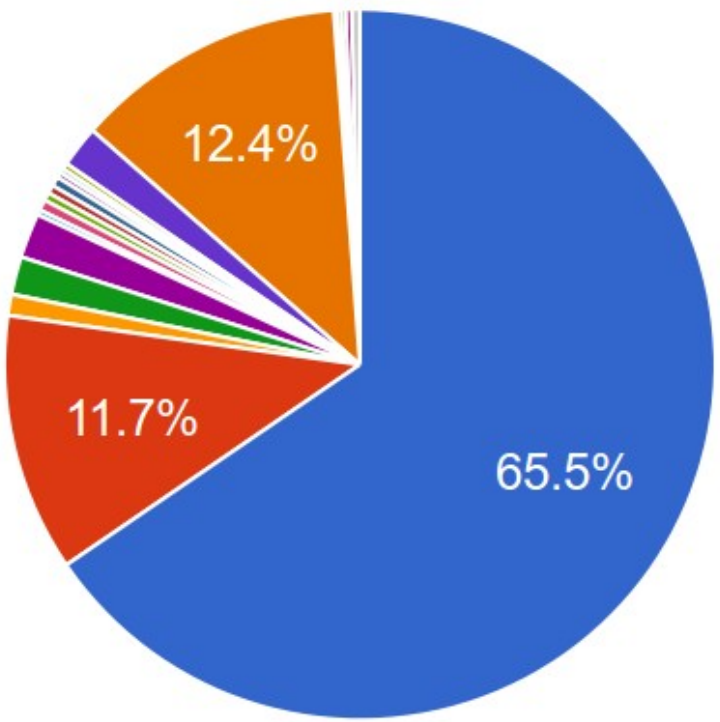
June 2014

Top500 Trends on Accelerated Architectures

Performance share:



June 2009



June 2014

Technology Path: Deeper Memory Hierarchy

High memory capability and capacity requirements

- Increasing compute performance
  ☞ **Increase of memory bandwidth B_{mem}**
- Applications ambition to solve large problems
  ☞ **Significant memory capacity C_{mem}**

Costs challenge

- Faster memory = more expensive (larger GByte/EUR)

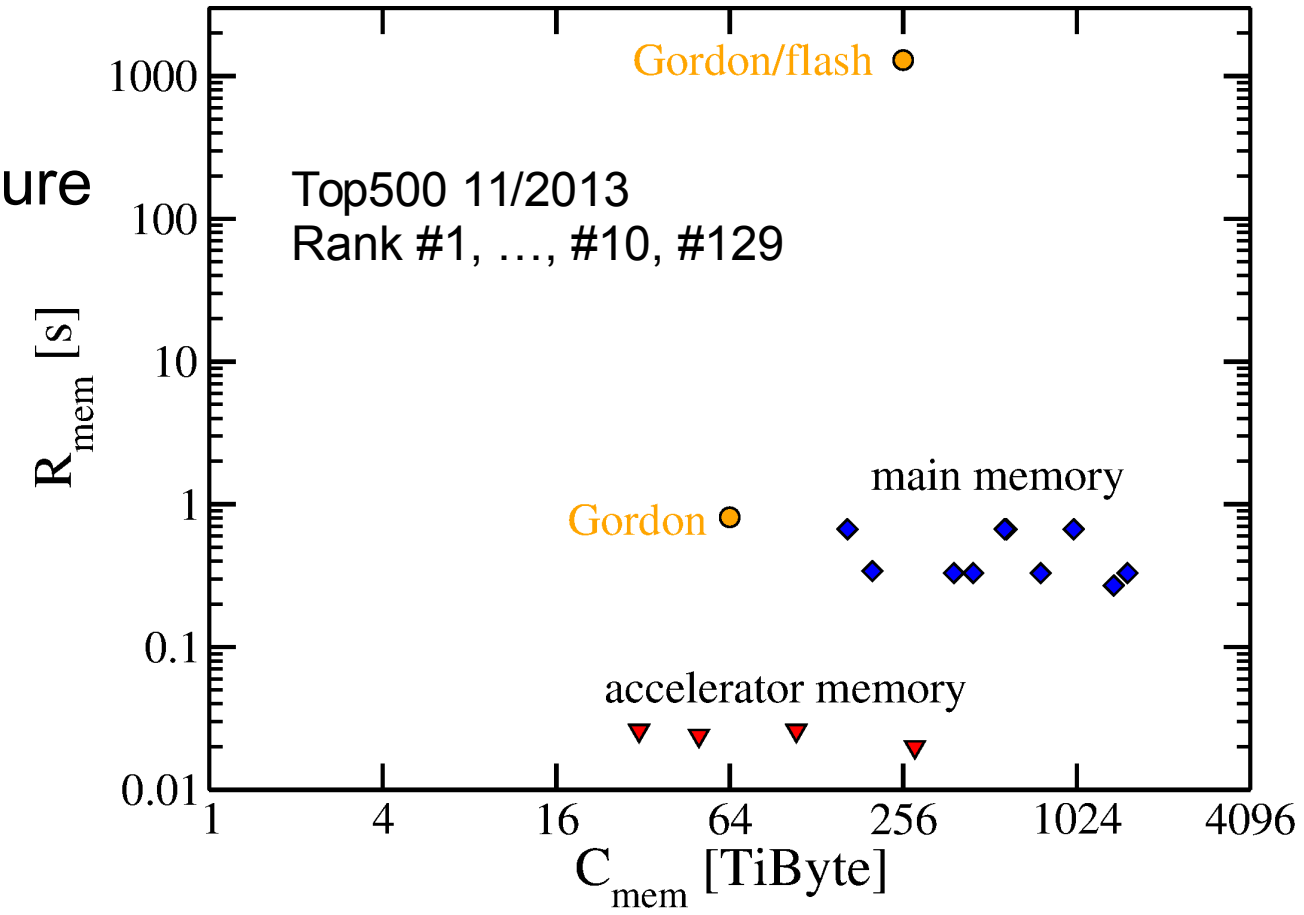
Solution: Memory hierarchy with more levels

- Fast memory, smaller capacity
- Large capacity, slower memory

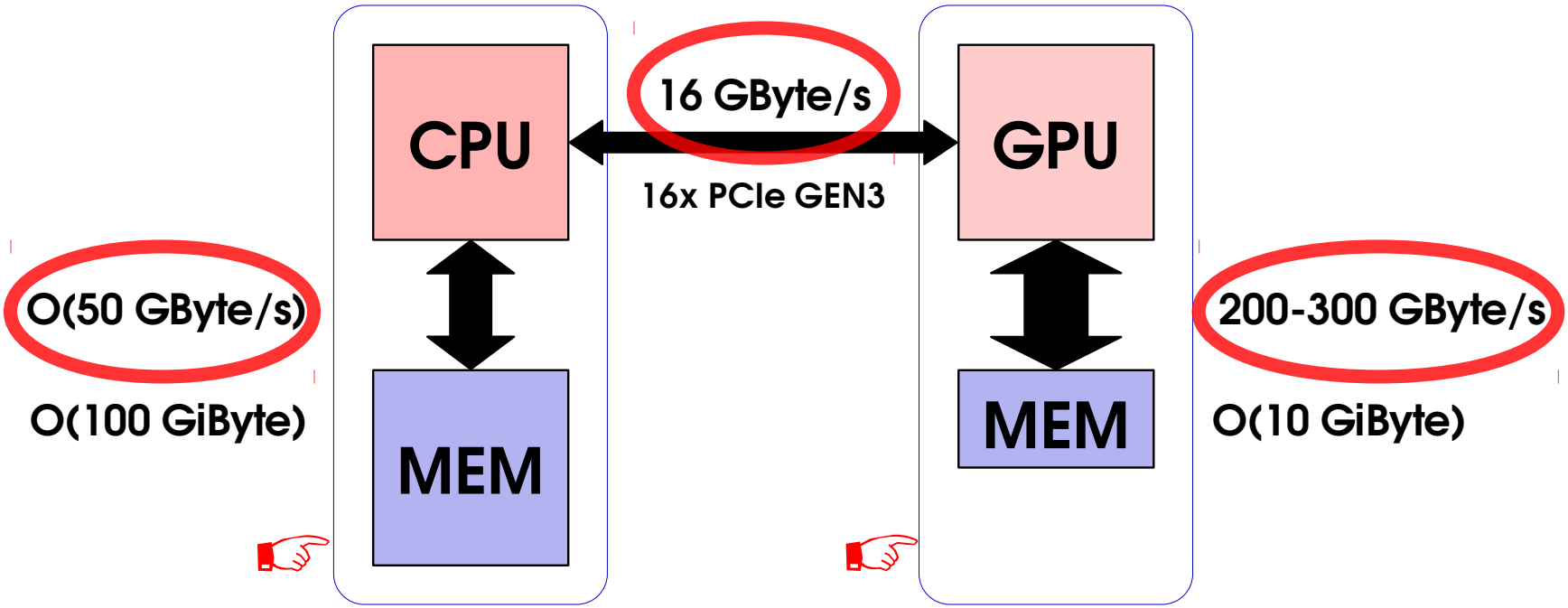
Deeper Memory Hierarchy: Top500 Trends

$R_{mem} = C_{mem}/B_{mem}$ vs. C_{mem}

- R_{mem} mainly determined by technology
- C_{mem} is architecture parameter

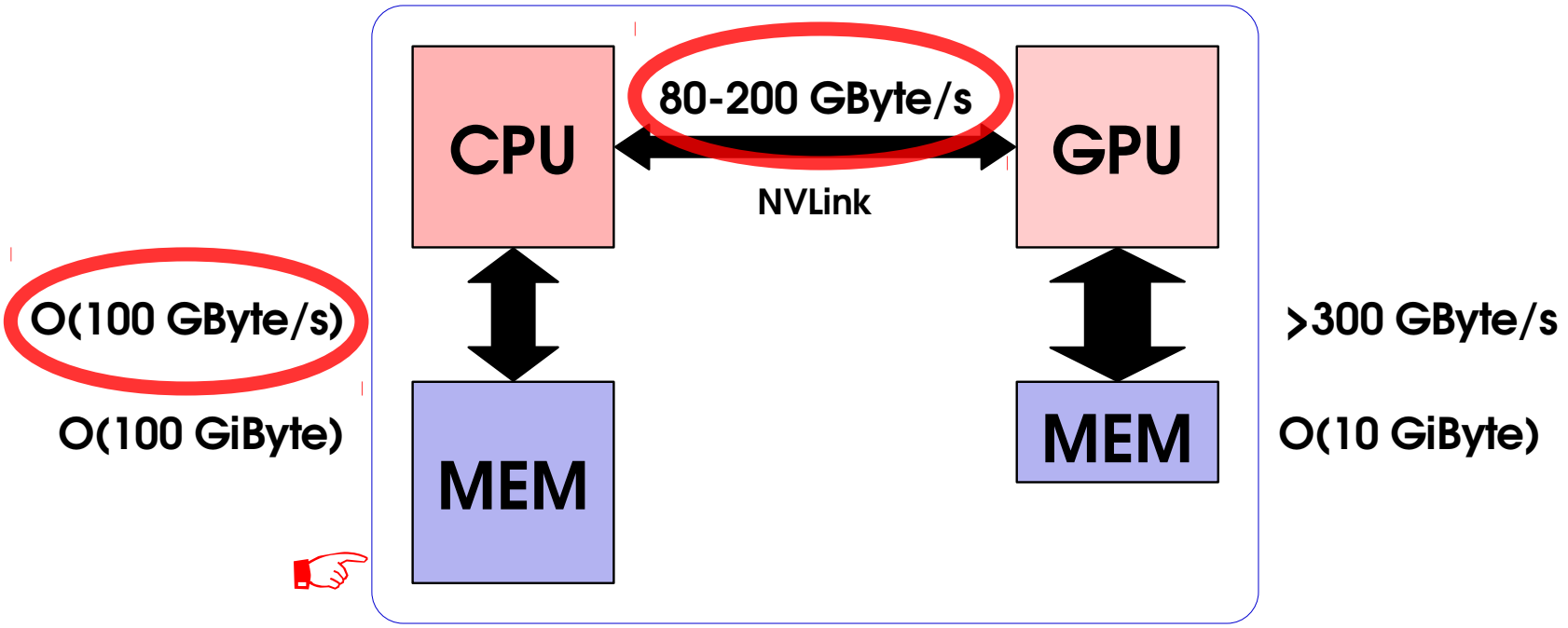


Accelerator Architectures Today



- Relatively small bandwidth between host and device
- Separate memory coherence domains

Future GPU Architectures



- Similar bandwidth host-device and host-memory
- Single memory coherence domains
- OpenPOWER is going down this road

OpenPOWER and the Exascale Challenges

Drastically improve energy efficiency

- GPU have potential for being highly energy efficient

Preserve usability at tremendously increased level of parallelism

- GPU architectures proven to be suitable for many scientific applications; growing experience and eco-system

Keep overall system balanced

- Tighter integration of CPU-GPU with different memory layers

Address reliability and resilience

- High-performance nodes → smaller number of components

Exascale Applications

Challenges for application developers

- Increase parallelism of application
- Manage data locality to leverage deeper memory hierarchy

☞ **Possible need for re-design of applications**

POWER Acceleration and Design Center

- Collaboration between IBM-BOE, IBM-ZRL, FZJ, NVIDIA
- Approach: Work on scalability of selected applications
- Create competence and knowledge for
  - Application developers
  - Technology developers

Summary and Conclusions

Need for more competition on HPC processors

- OpenPOWER provides solutions today
- Other alternatives are in the pipeline

Key technology trends towards exascale

- Massive increase of parallelism
- Deepening of the memory hierarchy

OpenPOWER drives architectures in right direction

- Tight integration of CPU and accelerator
- Improved usability of deeper memory hierarchy

Open eco-system good for co-design approach

☞ Good reasons for R&D along OpenPOWER roadmap